

DETECTION OF PHOTOVOLTAIC INSTALLATIONS IN RGB AERIAL IMAGING: A COMPARATIVE STUDY.

Steven Puttemans*, Wiebe Van Ranst and Toon Goedemé

EAVISE, KU Leuven - Campus De Nayer, Sint-Katelijne-Waver, Belgium
(steven.puttemans, wiebe.vanranst, toon.goedeme)@kuleuven.be

KEY WORDS: Photovoltaic Installations, Automated Object Detection, Fraud Detection

ABSTRACT:

In this work, we compare four different approaches for detecting photovoltaic installations from RGB aerial images. Our client, an electricity grid administrator, wants to hunt down fraud with unregistered illegal solar panel installations by detecting installations in aerial imagery and checking these against their database of registered installations. The detection of solar panels in these RGB images is a difficult task. Reasons are the relatively low resolution (at 25 cm/pixel an individual solar panel only measures about 9×7 pixels), the indiscriminating colour properties of the object (due to in-class variance and specular effects) and the apparent shape variability (rotation and skew due to the different roofs slant angles). Therefore, straightforward object segmentation techniques do not yield a satisfying solution, as proven in this paper. We compared four state-of-the-art object detection approaches for this task. First we experimented with a machine learning object detection technique based on pixel-based support vector machine classification. Secondly we developed an approach using MSER based colour segmentation and shape analysis. Finally a dual approach based on object categorization using the boosted cascade classifier technique of Viola & Jones and the aggregate channel features technique of Dollár et al., is introduced, learning a combination of colour and gradient feature based classifiers from a given training set. We successfully evaluate these four different approaches on a fully labelled test set of a 8000×8000 pixel, 4 square km zone containing 315 solar panel installations with in total more than 10.000 individual panels.

1. INTRODUCTION

Solar panels provide a solution in generating energy from non-polluting resources and therefore placing a solar panel installation is subsidized and encouraged by governments and electricity grid administrators. However due to the given funding and tax reduction, these 'greener' energy generating alternatives also give raise to malicious fraud. Our client, an electricity grid administrator wants to use the power of computer vision to track down these fraud cases to ensure that no financial benefits are given to people that are not correctly registered at the grid administrator with their solar panel installation.

Fully automated solar panel detection in RGB images yields some major challenges, as seen in Figure 1. First of all a solar panel is an object shape with only a few distinct visual features like shape and colour, due to its simple shape and colour distribution. Secondly the images that are freely available on the FGIA portal have a very limited resolution (25cm/pixel) resulting in a solar panel size of only 9×7 pixels, which is very little considering the real dimensions are around 1×1.5 m. This is clearly visible in Figure 1 (a). Due to the material properties of solar panels, specular reflections tend to change the visual properties of the panels, depending on the position of the sun. This could lead to solar pan-

els occurring as a flat white overexposed part of the image. This difference can be seen in Figure 1 (b) and (b'). The orientation in which the solar panels are placed also varies (south, south-east, south-west), raising the need of a full 360° rotational search of the image for object instances. Combined with the different roofs' slants (30° , 45° , 60° , etc.), panels get optically deformed and thus appear shorter or longer in 2D images, despite the fixed physical size, as seen in Figure 1 (c). Finally solar panels come in different materials (mono- and polycrystalline, full black, etc.) introducing even more intra class variance into the problem. For example, due to the small resolution per panel, full black installations appear as a black square, yielding no visual features to train computer vision techniques, as seen in Figure 1 (d).

Our goal is to develop a computer vision based approach that is able to automatically detect solar panels in aerial imagery and return the location of these installations with a high certainty. This in turn can avoid putting in huge amounts of manual labour to manually locate solar panel installations for fraud detection.

2. RELATED WORK

The automated analysis of solar panel installations from aerial images using techniques of the computer vision fields, limits itself to the analysis of solar panel efficiency and defects, using RGB and thermal cameras, as seen in (Sánchez-Friera et al., 2011; Li and Tsai, 2012; Tsai et al., 2013), while automated solar panel detection and localisation seem to be unexplored. While object detection is a well studied problem in the field of computer vision, many other fields still have not discovered the power of these state-of-the-art techniques in autonomous object detection and localisation. The community of computer vision has however already performed object detection research in the field of aerial imagery, focussing on roads (Hinz and Baumgartner, 2003), buildings (Mayer, 1999) or vehicles (Gleason et al., 2011), always using state-of-the-art computer vision algorithms and thus giving

*Corresponding author

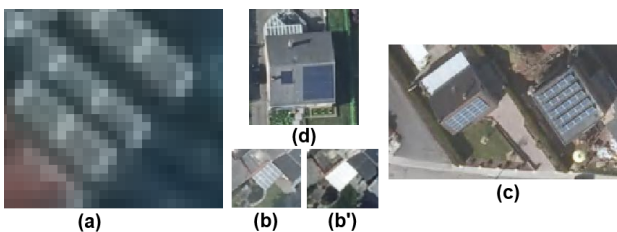


Figure 1. Challenges when dealing with solar panel detection.



Figure 2. Example of pixel based colour classification using neural networks as learning tool. (top) original image (middle) pixel classification result (bottom) cleaned up segmentation.

perfect example cases that can be expanded to solar panel detection and localisation.

While the intuitive way would be to use segmentation based approaches, where the RGB input image is transformed to a colour space where separating object pixels from background pixels using strict (learned) thresholds is easier, they tend to fail when a wide range of other objects in the images have similar colour ranges. Furthermore these techniques only take into account object colour information. Object categorisation techniques like (Viola and Jones, 2001; Felzenszwalb et al., 2010; Dollár et al., 2009; Dollár et al., 2014) tend to go one step further and use object properties like colour, shape, texture, etc. to uniquely define features that describe the object class, which are then used to successfully separate objects from the background.

3. DATASETS

For developing and testing our suggested approaches, we used the freely available medium-scale (25cm/pixel resolution) aerial footage from FGIA (Flanders Geographical Information Agency) covering the grid area of Flanders where our electricity grid administrator is active. From this publicly available dataset, a set of 2500 individual solar panels where manually annotated and a set of more than 150000 random negative samples (containing everything except solar panels) were collected and used to learn the models used by our state-of-the-art object detection algorithms.

To test the four suggested approaches an aerial image of 2×2 km of the city centre of Sint-Truiden was obtained. At a resolution of 25cm/pixel this results in a 8000×8000 pixel image, which is

up-scaled using a bi-cubic operator to 16000×16000 pixels to ensure the solar panels are covered by enough pixels per panel. Inside this test image 313 solar panel installations were manually annotated, by drawing polygons around the installations, to use as ground truth when validating the fully automated object detection techniques suggested in this work.

The complete dataset, including training data, test data and annotations can be found for research purposes at <http://www.eavise.be/SolarPanelDataset/>.

4. SUGGESTED APPROACH

To find an optimal solution for fully automated solar panel detection in aerial images, a comparative study was performed. In the following subsections each of the four state-of-the-art approaches, will be explained and discussed in detail. Section 5 will then take a closer look at the accuracy and time complexity achieved by each individual technique.

4.1 Pixel based colour classification using support vector machines

Our pixel based colour classification, as seen in Figure 2, uses the internal colour area of each solar panel (blue-grey colour range) without looking at the bright edges of the panel, in order to ensure that the pixel colour distribution of the training pixels are separable in the HSV colour space. We manually collected 1000 internal solar panel pixels and 2000 randomly selected non-solar panel background pixels. Both solar panel and non-solar panel pixels are transformed to the HSV colour space, a space where separating colours is easier than inside the RGB colour space. Based on this training data, a support vector machine classifier with linear kernel is trained, able to autonomously separate solar panel from non-solar panel pixels.

When a test image is presented to the support vector machine classifier, each pixel is processed and is given a certainty score, indicating how certain we are that the pixel is actually part of a solar panel installation. This certainty score allows us to set a minimal certainty threshold, generating a binary image as seen in Figure 2. On top of that, binary image opening and closing operators can be applied to remove noise, followed by contour detection and contour filling to achieve a cleaner result.

4.2 MSER based colour segmentation and shape analysis

Maximally stable extremal regions (Matas et al., 2004) is a technique used to detect blobs inside a given image. By systematically increasing the threshold on a given greyscale input image, from very sloppy to very strict, we create a set of sequential binary images. Inside those images the algorithm looks for regions that stay stable over the different thresholds and then approximates those regions by their fitted ellipse. Due to the higher response of solar panels in the blue channel compared to the red and green channels, we only process the blue channel data, removing the need to explicitly convert the RGB image to a greyscale image. Applying the MSER algorithm generates a selection of blob candidates, as seen in Figure 3(a). To further filter the obtained regions, we start by discarding blobs of an incorrect size (see Figure 3(b)), then look for blobs with a axis ratio that deviates maximally 30% of the ideal 1.5 : 1 ratio (see Figure 3(c)) and finally we apply HSV colour segmentation on the remaining blobs, using the technique discussed in subsection 4.1. The known size range of the blobs can be explained by the fact that aerial imagery is taken on a constant height, while the limited ratio deviation is explained by the fixed physical size of the solar panels. Finally

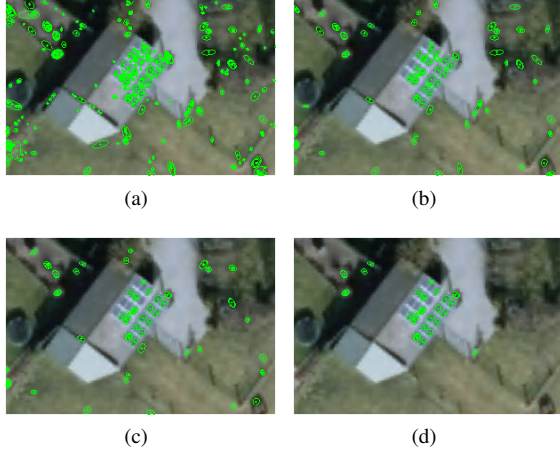


Figure 3. MSER based colour segmentation (a) MSER detected blobs (b) area restriction (c) ratio restriction (d) colour restriction.

since our solar panel installation training set contains solar panels with a known and shared colour range in the HSV colour space, we can allow colour based segmentation for further filtering.

4.3 Boosted cascade of weak classifiers

The previous techniques have only a very limited training time, since most processing is done on the fly when providing new test samples. This is different for object categorization techniques, where a model is learned from a set of positive object samples and a large set of random background samples. From each training sample a set of specific features is learned that is smartly combined into a model, able to separate objects from non-object patches inside the image.

For our first object categorization approach, we use the framework by Viola and Jones (2001), based on a boosted, using AdaBoost (Freund et al., 1999), cascade of weak classifiers (simple decision trees). This technique was originally developed for efficient face detection, but recent advances in computer vision (Puttemans et al., 2015; Puttemans and Goedemé, 2015) have proven that this technique still achieves top notch results in other application fields, focussing on more general object detection test cases.

The boosted cascade framework used does not incorporate colour information, but rather looks at the structure and texture of objects in a greyscale image. It describes local binary pattern features (Liao et al., 2007), features that look at local gradient information in training samples. Since solar panels have a colour range that

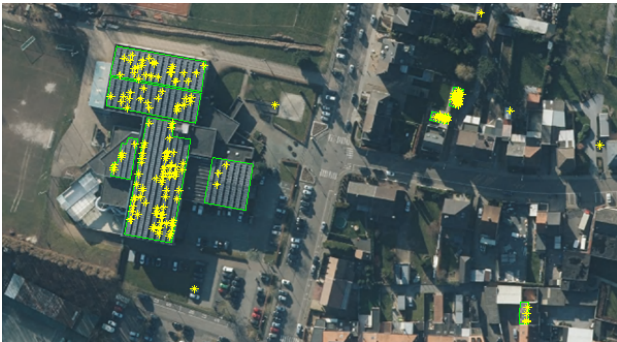


Figure 4. Detection of solar panels using a boosted cascade of weak classifiers.

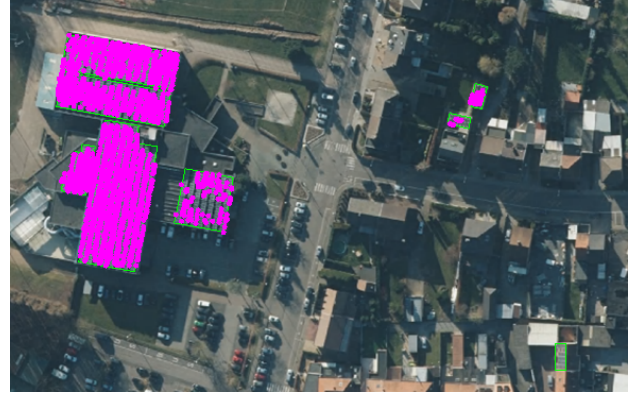


Figure 5. Detection of solar panels using the aggregate channel features technique.

has a higher response in the blue channel, we decided not to use a greyscale image, but the blue channel instead, explicitly forcing the framework to use colour information.

A downside of the boosted cascade classifier framework, is that it trains a model for a fixed orientation. For this we explicitly rotated all solar panel examples to a horizontal position, resulting in a model able to detect horizontal solar panels. However solar panels occur in different orientations. To cope with that, we simply rotate the input image over different angles, with a predefined angle step, and then warp back the retrieved detections. This allows us to build a full 360 degree capable solar panel detector using a single orientation model.

When running the detector on a test image, the detector will apply a fully rotational sliding window based evaluation of the image, triggering a detection at each position that gets classified as an object by the trained model. An example of such a detection output can be seen in Figure 4.

4.4 Aggregate channel features

The technique suggested by Dollár et al. (2009)(Dollár et al., 2014) is in fact an extension to the latter technique. Where the boosted cascade framework ignores colour information from the start, this technique proved that using colour information for certain object detection tasks can yield serious advantages and yield a higher detection accuracy. This is one of the main reasons why we decided to test this framework for our solar panel detection case. Besides that, we also have an internally developed C++ implementation of this framework (De Smedt and Goedemé, 2015) available, which allows us to easily run extra tests.

The technique does not take a single feature representation of the input image, like the previously used framework, but rather combines several feature channels, including colour, gradients, gradient histograms, etc. From this larger feature pool, the technique decides on its own which feature is good enough to efficiently separate positive and negative training data.

Running this more recent object categorization detector on top of a given test image generates similar output as the previous technique, as seen in Figure 5. However keep in mind that both Figure 4 and 5 are a sample detection output at specific detection thresholds of the algorithms. Deciding which algorithm performs better will be done thoroughly in section 5.

5. RESULTS

We started out with comparing our four state-of-the-art algorithms in processing time, as seen in Table 1. We can see a clear differ-

Table 1. Comparison of training (given the training sets described in section 3) and detection times (based on the upscaled 16000×16000 pixel image) combined with the complete system configurations used.

Technique used	Training time	Detection time	System configuration used
HSV + SVM	10 sec	10 sec	Intel(R) Core(TM) CPU i7-4500U @ 2.4GHZ
MSER	0 sec	100 sec	Intel(R) Core(TM) CPU i7-4500U @ 2.4GHZ
Boosted Cascade	3.5 hour	10 min	Intel(R) Xeon(R) CPU E5-2630 v2 @ 2.60GHz
Aggregate Channel Features	36 min	6 hour	Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40GHz

ence in training time between the more basic HSV pixel segmentation and the MSER approach on the one hand, and the object categorization techniques on the other hand. While the boosted cascade and the aggregate channel features approach take quite a bit longer to train on the given dataset, this task should only be done once, because a trained model can be reused as many times as we desire. However when looking at detection time, we notice a steady increase in processing time when computational complexity of the algorithm increases. Where the standard pixel based segmentation takes only 10 seconds for a 16000×16000 pixel image, the basic object categorization framework already takes 600 times that long.

However these timings should be interpreted with caution because they highly depend on the available infrastructure, which is also specified in Table 1. Furthermore the implementation of the aggregate channel features technique was developed in-house and was not yet optimized for parallel processing, thus needing to process everything in a sequential order.

One of the main reasons why object categorization techniques take a lot more time, is because these techniques are trained for a specific orientation. In order to be able to detect objects in every possible orientation, we rotated the original image for a full 360 degrees, with a single degree step. This again can be heavily optimized when desired, by for example rotating the model instead of rotating the image.

To evaluate the accuracy of the developed techniques, we suggest to use precision recall curves, used to compare the actual detection output with the manually obtained ground truth polygons. To generate these curves, seen in Figure 6, the generated detection maps are first downsampled to a 4000×4000 pixels, combining scores of the detections obtained on the same locations. On top of the resulting score map, a threshold is applied (which is the varying parameter used to generate the different precision-recall values for each algorithm). This is followed by a dilation (make detection centres as large as solar panel dimensions) and erosion

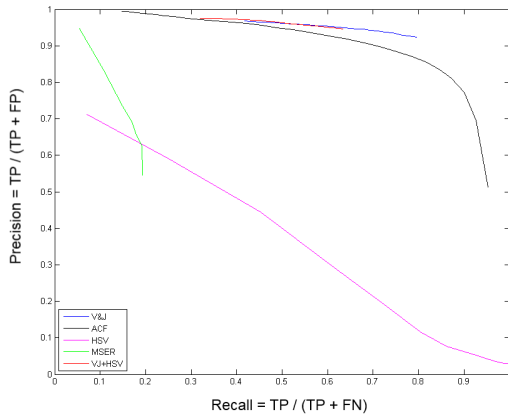


Figure 6. Precision - Recall curves for all techniques tested and validated on the 16000×16000 test image of the Sint-Truiden city centre.

(remove detections that are lonely and not grouped) operations resulting in a clear binary image. This binary image is then compared to calculate the amount of true positives, false positives and false negatives, which are in turn used to calculate correct precision and recall values.

Figure 7 shows the comparison of that binary output which is in turn compared to the manual annotations. By using the three channels of an RGB image, knowledge of both results can be visually combined. Detection centres are visualised as dots on the red channel, the ground truth is visualized on the green channel and finally the detection output regions are visualised on the blue channel. Combining those three colour channels yields a set of pixel based classification labels. Cyan labels indicate true positive detections, black labels indicate true negative detections, blue labels indicate false positive detections and green labels indicate false negative detections. The visualisation is done for both the boosted cascade and the aggregate channel technique.

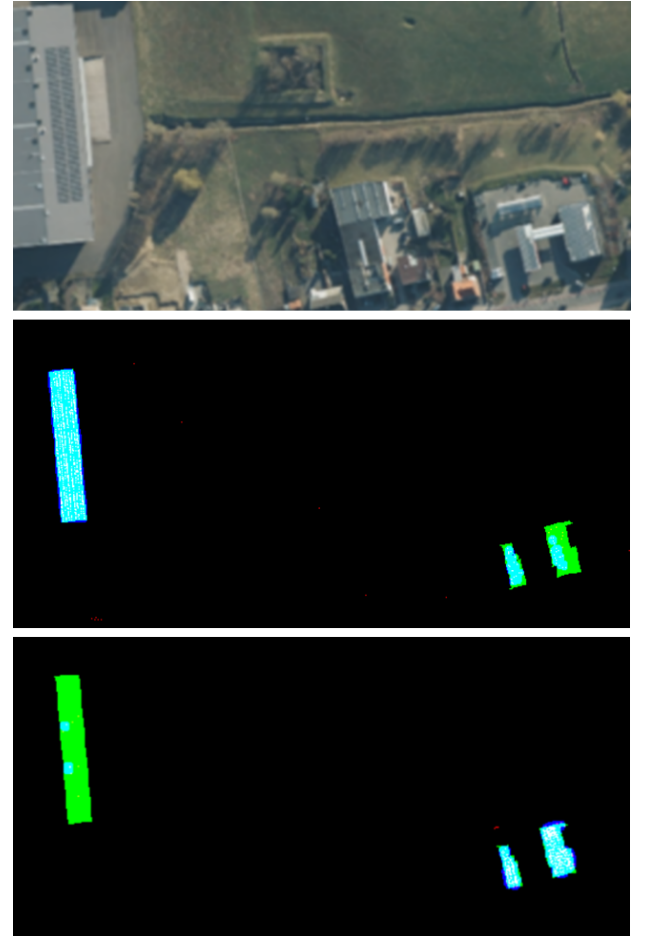


Figure 7. Score processing map for both boosted cascade and aggregate channel features technique. (top) original input image (middle) boosted cascade result (bottom) aggregate channel features result

The precision-recall curves clearly show that object categorization techniques outperform the other, more basic computer vision techniques. We notice that the boosted cascade technique performs still a bit better than the aggregate channel technique, which indicates that adding extra feature channels like colour and other gradient filters is overkill for solar panel detection. To test this theory we added HSV pixel based segmentation as a post-processing step to our boosted cascade detector, displayed as the red curve in Figure 6. We see almost no increase in efficiency which made us to decide to drop this extra processing step, in order to obtain a smaller computational complexity and thus a faster running time.

6. DISCUSSION AND CONCLUSION

Our goal is to compare techniques for automatically detecting photovoltaic solar panel installations in RGB aerial images, considering the very challenging conditions due to the limited resolution, the limited amount of visual features, the existence of specular reflections and the different orientations of the solar panel object class.

Using pixel based segmentation for solar panel detection yields only moderate results. In any given test image, it is impossible to only separate blobs that are pure solar panels, since the colour distribution of the panels, also tends to return in many background patches. Furthermore there are several solar panel types that tend to deviate from the average colour distribution, like the full black type, generating even more missed detections.

Applying smarter feature based techniques, like the MSER based approach, seems promising and even yields decent results in specific areas, but still has quite a fast drop in accuracy when considering larger regions where solar panels need to be detected, as seen from the resulting precision-recall curve based on the 2×2 km area of Sint-Truiden.

However, our tests clearly showed that object categorization techniques like boosted cascades and aggregate channel features can outperform more basic techniques with the only downside that training data needs to be collected and that the single training setup for building the model takes a bit longer.

We proved that computer vision is a working solution for efficient and fully automated solar panel detection in RGB based aerial imagery for fraud detection. For a fully automated system, achieving a precision of 93% at a recall of 80% is not bad at all. Certainly if you take into account some considerations. Due to the heavy deformations a roofs' slant angle can introduce to the solar panel object, training multiple models to cope with the missed detections could be a first improvement. Furthermore, we performed the research on mid resolution images, at 25cm/pixel. However there are already measurements of the same area at a resolution of 8cm/pixel available, which would again yield an accuracy increase. Since our training set consists mainly of industrial solar panel arrays, it is also quite understandable that the model is not able to cope with the larger deformations found on solar panels placed on domestic housing.

Finally one could argue that the fully automated system should not be able to detect 100% of the solar panels with 100% precision. In most set-ups, an operator is still available for slight semi-automated input. We could provide that percentage of object that we did not found, or that yield a lower certainty, to the operator and ask him to define these cases of doubt by manual inspection. This suggested semi-automated approach would still reduce the manual labour drastically and reduce the time needed to process larger datasets.

As future research based on this paper we suggest to take a closer look at CNN based detectors Girshick et al. (2014). In many object detection tasks they achieve top notch performance and could thus improve the obtained results.

7. ACKNOWLEDGEMENTS

This work is supported by the KU Leuven, Campus De Nayer, through a scholarship. We would like to thank Infrax, for providing the solar panel datasets and manual annotations.

References

- De Smedt, F. and Goedemé, T., 2015. Open framework for combined pedestrian detection. In: *VISAPP*, Vol. 2, pp. 551–559.
- Dollár, P., Appel, R. et al., 2014. Fast feature pyramids for object detection. *TPAMI* 36(8), pp. 1532–1545.
- Dollár, P., Tu, Z. et al., 2009. Integral channel features. In: *BMVC*.
- Felzenszwalb, P. F., Girshick, R. B. et al., 2010. Cascade object detection with deformable part models. In: *CVPR*, pp. 2241–2248.
- Freund, Y., Schapire, R. et al., 1999. A short introduction to boosting. *JSACI Journal* 14(771–780), pp. 1612.
- Girshick, R., Donahue, J. et al., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In: *CVPR*, pp. 580–587.
- Gleason, J., Nefian, A. et al., 2011. Vehicle detection from aerial imagery. In: *ICRA*, IEEE, pp. 2065–2070.
- Hinz, S. and Baumgartner, A., 2003. Automatic extraction of urban road networks from multi-view aerial imagery. *Journal of Photogrammetry and Remote Sensing* 58(1), pp. 83–98.
- Li, W.-C. and Tsai, D.-M., 2012. Wavelet-based defect detection in solar wafer images with inhomogeneous texture. *Pattern Recognition* 45(2), pp. 742–756.
- Liao, S., Zhu, X. et al., 2007. Learning multi-scale block local binary patterns for face recognition. In: *Advances in Biometrics*, Springer, pp. 828–837.
- Matas, J., Chum, O. et al., 2004. Robust wide-baseline stereo from maximally stable extremal regions. *Image and vision computing* 22(10), pp. 761–767.
- Mayer, H., 1999. Automatic object extraction from aerial imagery a survey focusing on buildings. *CVIU* 74(2), pp. 138–149.
- Puttemans, S. and Goedemé, T., 2015. Visual detection and species classification of orchid flowers. In: *MVA*, IEEE, pp. 505–509.
- Puttemans, S., Goedemé, T. et al., 2015. Automated walking aid detector based on indoor video recordings. In: *EMBC*, IEEE, pp. 5040–5045.
- Sánchez-Friera, P., Piliouguine, M. et al., 2011. Analysis of degradation mechanisms of crystalline silicon pv modules after 12 years of operation in southern europe. *Progress in photovoltaics: Research and Applications* 19(6), pp. 658–666.
- Tsai, D.-M., Wu, S.-C. et al., 2013. Defect detection in solar modules using ica basis images. *IEEE Transactions on Industrial Informatics* 9(1), pp. 122–131.
- Viola, P. and Jones, M., 2001. Rapid object detection using a boosted cascade of simple features. In: *CVPR*, Vol. 1, IEEE, pp. I–511.